

A propósito del COVID19, una revisión sistemática respecto al paradigma de la radiología y la inteligencia artificial

Gabriela Puentes Suarez¹  Harvey Segundo Valencia Gamarra²  Carlos Gonzalo Arteaga Rodriguez³  Yohan Lilibeth Suárez Uribe⁴  Héctor Javier Espitia Soto⁵ 

1 Gabriela Puentes Suarez*: Universidad del Rosario; puentes.gabriela@gmail.com

2 Harvey Segundo Valencia Gamarra; Fundación Universitaria Juan N. Corpas; md.harveyvalencia@yahoo.com.co

3 Carlos Gonzalo Arteaga Rodriguez; Universidad Tecnológica de Pereira; cgarateaga@utp.edu.co

4 Yohan Lilibeth Suárez Uribe, Fundación Universitaria Juan N. Corpas; johansuarez1@gmail.com

5 Héctor Javier Espitia Soto, Fundación Universitaria Sanitas, drawco8720@hotmail.com

Historia del Artículo:

Recibido el 10 de febrero de 2021

Aceptado el 3 de abril de 2021

On-line el 15 de abril de 2021

Palabras Clave:

Radiología, Pulmón, Inteligencia Artificial, Comparación, Neumonía.

Keywords:

Radiology, Lung, Artificial Intelligence, Comparison, Pneumonia.

Resumen

La inteligencia artificial (IA) es un campo de rápido crecimiento que plantea la posible solución a múltiples problemas presentados, el día de hoy, en la práctica clínica, particularmente, en el área de radiología. En principio se puede entender a la IA como la habilidad de una computadora para imitar la capacidad cognitiva humana; corresponde a una serie compleja de actividades, entre ellas, el aprendizaje automático y los algoritmos de procesamiento, estos se encuentran basados en parámetros de conocimiento predefinido que pueden procesar grandes volúmenes de información imperceptibles para el ojo humano, lo que abre una nueva era para el desarrollo de la tecnología de imagen diagnóstica.

Abstract

Artificial intelligence is a rapidly growing field that poses the possible solution to multiple problems arising today in clinical practice, particularly in the area of radiology. In principle, Artificial Intelligence (AI) can be understood as the ability of a computer to imitate human cognitive capacity, it corresponds to a complex series of activities including machine learning and processing algorithms which are based on predefined knowledge parameters that can process large volumes of information imperceptible to the human eye, opening a new era for the development of diagnostic imaging technology.

* Autor para correspondencia:

Gabriela Puentes Suarez; Universidad del Rosario, e-mail: puentes.gabriela@gmail.com

Cómo citar:

Puentes et al. A propósito del COVID19, una revisión sistemática respecto al paradigma de la radiología y la inteligencia artificial. S&EMJ. Año 2021; Vol. 2: 142-151.

Objetivo

Determinar si existe una diferencia, en términos de rendimiento diagnóstico, en pacientes con alteraciones pulmonares al aplicar herramientas de IA en comparación a técnicas de imagen convencional.

Método

Se realiza una búsqueda sistemática con los términos Mesh: *pneumonia, radiology, artificial intelligence, comparison* y *lung*, junto con los operadores booleanos AND, OR, en las bases de datos PubMed, Ovid, Cochrane, BMJ y LILACS, con restricción de lenguaje para inglés y español. Los criterios de elegibilidad atañen a estudios de cohortes de radiología e IA.

Conclusiones

La pandemia por COVID-19 es un escenario que ha permitido impulsar la implementación de la IA en la práctica clínica, esto fundamentado en la falta de personal, los recursos limitados y la necesidad de herramientas más efectivas para realizar diagnósticos. Pese a que existen varios estudios que describen su desarrollo y posibles aplicaciones, pocos validan estos modelos. Sin embargo, los estudios publicados, hasta el momento, han demostrado un mayor rendimiento diagnóstico del uso de IA frente a su no uso, particularmente para el diagnóstico de neumonía por COVID-19, pues tienen un potencial para mejorar la toma de decisiones dada la alta especificidad y sensibilidad diagnósticas.

COVID -19, a systematic review regarding the paradigm of radiology and artificial intelligence

Objective

To identify the difference between using artificial intelligence and traditional imaging in patients with pulmonary pathologies.

Method

A systematic search was carried out with the Mesh terms: Pneumonia, Radiology, Intelligent Artificial, AI artificial intelligence, Comparison and Lung, together with the Boolean operators AND, OR, in the PubMed, Ovid, Cochrane, BMJ and LILACS databases, with language restriction to English and Spanish. Of the total results, 400 abstracts were reviewed, and 50 complete articles were selected. The eligibility criteria corresponded to radiology and artificial intelligence cohort studies, and studies were chosen that will compare the AI system evaluated with expert radiologists, or in which there was no comparison. The non-COVID-19 pneumonia group includes other types of pneumonia, which are not caused by COVID-19, such as viral or bacterial pneumonia, as well

as influenza A and SARS. Other studies such as chest radiography and scintigraphy were excluded. Likewise, articles evaluating COVID-19 severity were not taken into account.

Conclusions

The diagnosis of COVID-19 pneumonia using AI algorithms has great potential. In this sense, it is pointed out with some caution that, AI algorithms can be used for diagnosis in COVID-19 using medical radiology images with an equivalent or even better performance for health professionals, in terms of sensitivity and specificity, providing a basis for its clinical application. Similarly, it is thought that the clinical application of AI can alleviate the shortage of medical resources, improve the detection rate and precision and then the prognosis of patients.

Los modelos de IA se formalizaron en los años 60, inicialmente se desarrollaron dos tipos: el cognitivismo, entendido como el impulso de programas basados en reglas, y el conexionismo, desarrollo de programas ingenuos que operaban a través de datos secundarios (1). Los modelos de IA, usados en la actualidad, atañen al conexionismo, y uno de los más conocidos es el de neuronas artificiales, estas se organizan en redes artificiales que permiten que estos programas "puedan ser educados" (2,3). Los modelos más tradicionales se sustentan en características y etiquetas predefinidas, estos parámetros permiten cuantificar particularidades radiográficas específicas tales como la forma, la textura, la distribución y las densidades de los píxeles (2); la información recopilada se procesa mediante modelos estadísticos de aprendizaje que se ajustan a los datos para identificar biomarcadores en imágenes (3). Estas redes son capaces de realizar, automáticamente, un reconocimiento de imagen, donde cada dato recogido es asociado con una etiqueta, como es el ejemplo de un diagnóstico predefinido (1).

A este tipo de modelo, basado en redes neuronales, se le conoce también como aprendizaje profundo, esta es una subrama del aprendizaje automático (4). Los programas sobre este último se asientan en el análisis de imágenes con la menor intervención humana posible, esto con el fin de posibilitar que los algoritmos decidan cuáles son los hallazgos más relevantes de cada caso (5).

Cabe destacar que, con el paso de los años y con el creciente desarrollo de la tecnología, particularmente el análisis de Bigdata, la IA tomó un papel cada más relevante para la medicina, donde tuvo su mayor proyección en el 2010 debido al progreso de las redes neuronales (6,7). No obstante, la aplicación real en la práctica clínica se ha encontrado con diversas limitaciones y su aplicación ha generado múltiples detractores, debates e incógnitas (8). Aunque por el momento faltan datos suficientes para la aplicación rutinaria de esta tecnología en la práctica diaria, lo

más lógico es suponer que el desarrollo de soluciones automatizadas, apoyadas en el aprendizaje profundo, comenzará a abordar los problemas clínicos más comunes cuando haya suficientes datos disponibles (5). En este contexto, la pandemia permite probar las aplicaciones y limitaciones de IA en el proceso de nutrir el debate que rodea a su implementación (1,6,9).

En este orden de ideas, existen dos métodos de IA para una tarea de clasificación representativa. El primero se basa en características de ingeniería extraídas de regiones de interés sobre la base del conocimiento de expertos, estas singularidades se seleccionan y se incorporan a los clasificadores de aprendizaje automático (1). El segundo utiliza un aprendizaje profundo y no requiere anotación de región, debido a que la localización suele ser suficiente; comprende varias capas donde la extracción de características, la selección y la clasificación final se efectúan simultáneamente durante el entrenamiento (10). A medida que las capas aprenden características de nivel cada vez más alto, las capas anteriores pueden asimilar formas abstractas mientras que otras más profundas pueden captar objetos completos (10) (Figura 1).

Figura 1. Esquema del aprendizaje profundo: cuando las capas retienen características de nivel cada vez más alto, las anteriores pueden asimilar líneas y sombras, y otras más profundas pueden aprender órganos. En las fases delineadas de color naranja se esquematizan las capas de convolución para la extracción de características. Las azules sintetizan la agrupación de capas para la agregación de particulares y las verdes la conexión de estas (11).



Fuente: Adaptada para fines académicos de: *Artificial intelligence in radiology*. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. *Artificial intelligence in radiology*. (11)

Los datos son el componente central y crucial para el aprendizaje de los sistemas de IA (11,13). Está claro que existen grandes cantidades de datos médicos y se almacenan para permitir una facilidad moderada de acceso y recuperación (1,14), empero, rara vez se usan para facilitar el aprendizaje de cualquier modelo de IA (14,15). El desempeño subóptimo de muchos algoritmos de segmentación automatizados y semiautomatizados ha obstaculizado su utilidad en el uso de datos, en virtud de que casi siempre se necesitan lectores humanos para verificar la precisión (14,16). De este modo, una solución que viabiliza su conservación automatizada es el aprendizaje no supervisado (17); los avances recientes en este son muy prometedores, pues las características discriminatorias se aprenden sin un etiquetado explícito (14,17).

Los avances actuales en el hardware de imágenes, en términos de calidad, sensibilidad y resolución,

consienten la discriminación de diferencias mínimas en los tejidos (18,19). Estas disimilitudes son difíciles de reconocer por un ojo entrenado y por algunos métodos tradicionales de IA empleados en la clínica (20); las mejoras prometen una mayor precisión y las funciones de los radiólogos se amplían cuando se conectan más a la tecnología y tienen acceso a mejores herramientas (20,21). También es probable que surjan como elementos críticos en el proceso de entrenamiento de IA, con el fin de aportar conocimiento y supervisar la eficacia (21).

Dado que las diferentes modalidades de IA superan el rendimiento humano, se espera que se convierta en un recurso educativo valioso (17). Los operadores humanos supervisan los resultados y buscan interpretar el razonamiento detrás de ellos, esto como un medio de validación y un recurso para descubrir información oculta que pudo haber pasado por alto (17).

COVID-19 y tecnologías usuales

El COVID-19 se diagnostica con pruebas moleculares y de anticuerpos, a pesar de esto, su disponibilidad limitada ha dificultado la detección rápida y la contención de los pacientes con el virus (22). En este punto, la radiología tiene un papel crucial en el diagnóstico de pacientes con sospecha de COVID-19 (23,24).

La radiografía de tórax es económica y accesible (25), es un método de fácil consecución que posibilita la orientación del paciente al ingreso, en especial aquellos sospechosos de poseer el virus y con características clínicas compatibles con infección moderada o severa (26). Este abordaje inicial ayuda a detectar particularidades relacionadas con neumonía (27), asimismo, consiente proporcionar un diagnóstico alternativo. Pero una radiografía simple de tórax no puede excluir el diagnóstico de COVID-19, debido a que su sensibilidad depende del momento de la obtención de imágenes y la gravedad de la afectación pulmonar (28).

Otro amplio uso de la radiografía se da en las salas de emergencia y cuidado intensivo, estas son portátiles y fáciles de emplear, esto favorece el seguimiento de la gravedad y la progresión de la enfermedad sin necesidad de movilizar al paciente (26). La radiografía de tórax suele ser muy útil en el acompañamiento de pacientes hospitalizados, especialmente para la evaluación de complicaciones como neumotórax y derrame pleural (25).

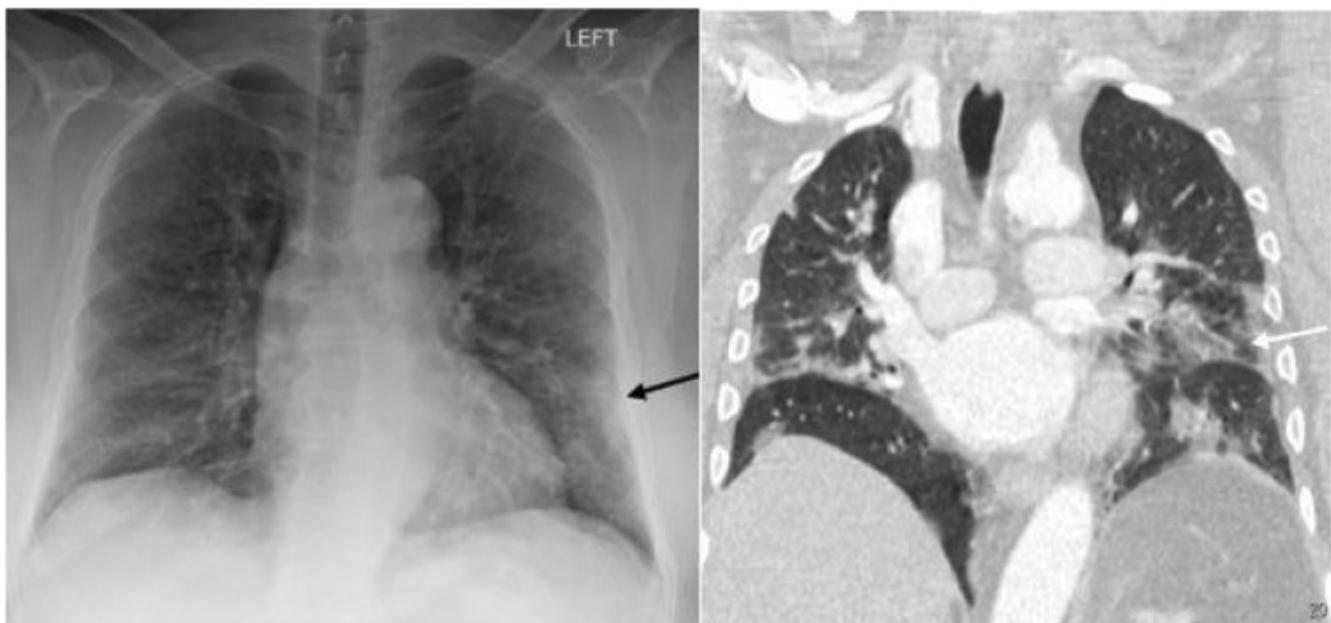
La radiografía de tórax es una modalidad menos sensible en la detección de la enfermedad pulmonar por COVID-19, en comparación con la tomografía computarizada (TC) (3). Los hallazgos de COVID-19 más comunes incluyen la consolidación pulmonar y las opacidades en vidrio esmerilado (3). No obstante, con una sensibilidad aproximada del 69 %, las densidades de vidrio esmerilado, observadas en la tomografía, a menudo son extremadamente difíciles de correlacionar (29). Las opacidades pulmonares a veces pueden ser

difusas, lo que dificulta su identificación (Figura 2) (29).

Otro descubrimiento para considerar en la radiografía de tórax es la neumonía bilateral (26). A diferencia de la neumonía bacteriana adquirida en la comunidad, el COVID-19 produce opacidades pulmonares en más de un lóbulo y frecuentemente de forma bilateral (26,29). La identificación de la enfermedad multifocal del espacio aéreo en radiografía puede sugerir la neumonía de dicho virus (27); la enfermedad del espacio aéreo tiende a tener una distribución pulmonar baja (25). Igualmente, una de las características más específicas de la neumonía de Covid-19 es la alta frecuencia de afectación pulmonar periférica, esta refleja, habitualmente, otros procesos inflamatorios como la neumonía organizada (26). Dichas opacidades pulmonares periféricas también tienden a ser multifocales, en parches o confluentes, y pueden identificarse fácilmente en la radiografía de tórax (29).

En tal marco, un evento relevante es la enfermedad difusa del espacio aéreo. Las opacidades pulmonares difusas en pacientes con COVID-19 tienen un patrón similar con otros procesos infecciosos o inflamatorios generalizados (25). Estas pueden evolucionar rápidamente a un patrón de consolidación dentro de 1 a 3 semanas de la aparición de los síntomas, con frecuencia alcanzan su punto máximo alrededor de los 6 a 12 días después de la presentación clínica inicial (25) many physicians in a wide variety of specialties continue to play pivotal roles in diagnosis and management. In radiology, much of the literature to date has focused on chest CT manifestations of COVID-19 (Zhou et al. [1]; Chung et al. [2]).

Figura 2. Radiografía de tórax con opacidades en vidrio esmerilado en el lóbulo medio e inferior izquierdo periférico en parches (flecha negra), correspondientes estos a opacidades en vidrio esmerilado (flecha blanca) en la imagen tomográfica.



Fuente: Tomada para fines académicos de: *portable chest X-ray in coronavirus disease-19 (COVID-19): a pictorial review* (25).

La TC de tórax proporciona una delimitación con una alta sensibilidad de hasta el 98 % en neumonía por COVID-19 (30). Empero, el uso de la TC como prueba de diagnóstico para el virus se ha cuestionado, pues no otorga ningún beneficio clínico adicional, pero podría dar lugar a una falsa sensación de seguridad, en vista de que el 20 % de los pacientes sintomáticos tienen resultados negativos en la TC (31). El resultado de una TC de tórax no altera el manejo del paciente, múltiples sociedades recomiendan no utilizarla para el cribado y el diagnóstico de la enfermedad (31).

Los hallazgos típicos de la TC incluyen la distribución bilateral de opacidades en vidrio esmerilado en los pulmones periféricos y posteriores (32,33). Cada vez que avanza la enfermedad, las opacidades en vidrio esmerilado pueden aumentar en tamaño y en grado de afectación, en este sentido, se pueden observar patrones de pavimentación o consolidaciones adicionales (31, 32). La presencia de consolidaciones con broncograma aéreo, afectación pulmonar central y derrame pleural en la TC de tórax inicial se presenta, con mayor frecuencia, en pacientes graves que necesitan cuidados intensivos (33).

Covid 19 e inteligencia artificial

Se han empleado algoritmos de IA para ayudar a los radiólogos a interpretar imágenes con mayor rapidez y precisión en esta pandemia. Un estudio temprano mostró que la IA puede acrecentar el desempeño de los radiólogos para distinguir el COVID-19 de la neumonía de otras etiologías en la TC de tórax, lo que arroja una mayor precisión (90 %), sensibilidad (88 %) y especificidad (96 %) (5,34).

Mediante el análisis de la radiómica de la TC y los

factores clínicos y demográficos, los investigadores han desarrollado modelos de aprendizaje automático que pueden predecir la probabilidad de los pacientes con COVID-19 en requerir ventilación mecánica con una precisión prometedora del 75 % (15,35). Un modelo de aprendizaje profundo, ejecutado recientemente, tiene la posibilidad de suministrar una solución que pueda identificar el virus en las radiografías de tórax con mayor precisión que los radiólogos, con un área bajo la curva característica de funcionamiento del receptor de 0.81 %, sensibilidad del 84.7 % y especificidad del 71.6 % (36).

Se han creado algoritmos para apoyar el procedimiento general de aplicaciones basadas en IA que ayuden a los médicos a tomar una ruta en el paciente con COVID-19 (34).

La Figura 3 es una adaptación del algoritmo creado por Vaisha y colaboradores, el diagrama de flujo de la Figura 5, explica cómo algunos estudios sugieren que la participación de la IA, en los pasos importantes del tratamiento de alta precisión, reduce la complejidad y el tiempo necesario (34). Así, el médico no solo se centra en el tratamiento del paciente, sino en el control de la enfermedad con la aplicación de IA (34).

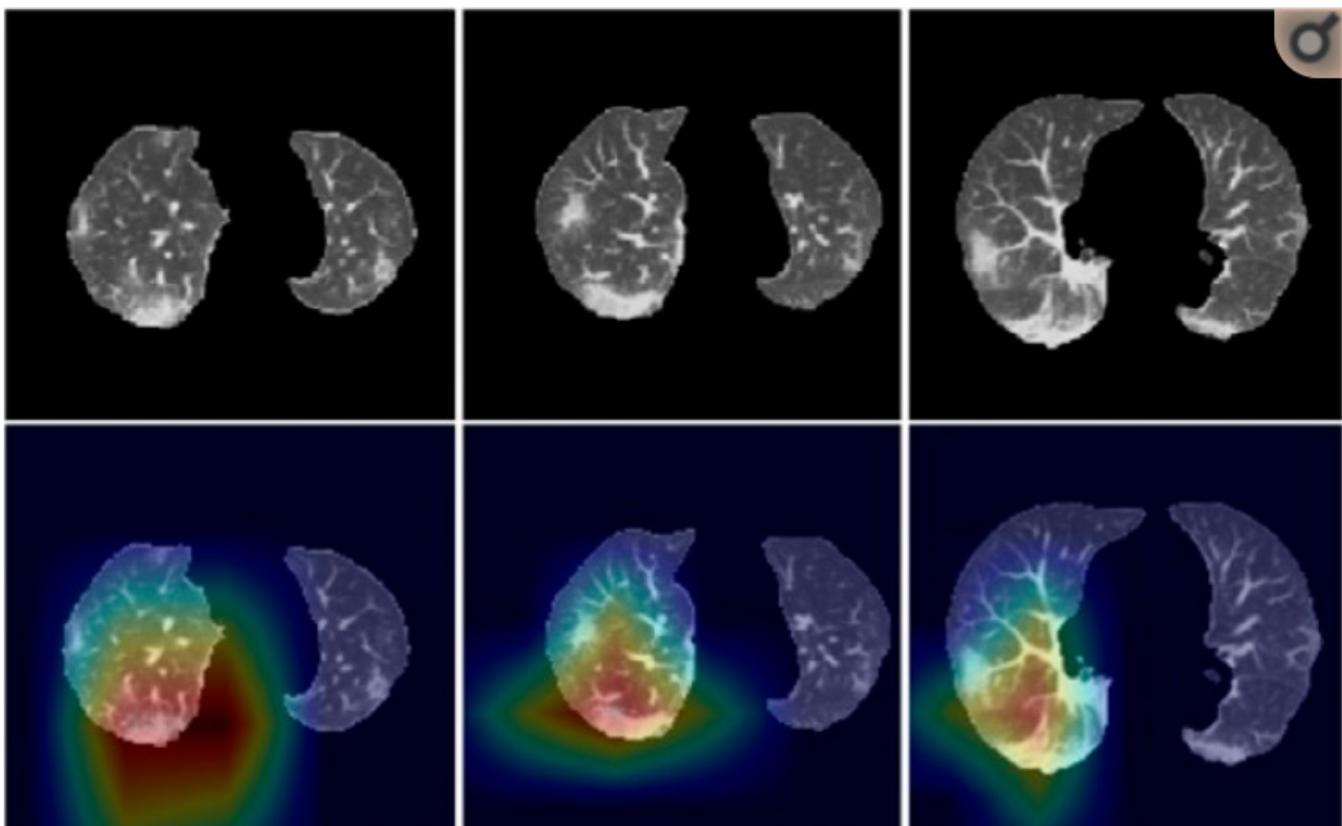
Algunos autores argumentaron las ventajas de la

IA en el diagnóstico y manejo del COVID-19. El primero es un diagnóstico precoz, en virtud de que se ha propuesto que la IA puede cooperar para promover un nuevo sistema de diagnóstico y gestión de los casos de COVID 19, esto por medio de algoritmos útiles (8). La IA es funcional en el análisis de los casos infectados con la ayuda de tecnologías de imágenes médicas como la TC (37), adicionalmente, se utiliza para reducir la carga de trabajo del personal de la salud (7). Igualmente, esta posibilidad efectuar un diagnóstico temprano y brindar tratamiento en una etapa anticipada con el uso de enfoques digitales y ciencia de decisión, y con el propósito de ofrecer la mejor capacitación a estudiantes y médicos con respecto a esta nueva enfermedad (7).

Resultados

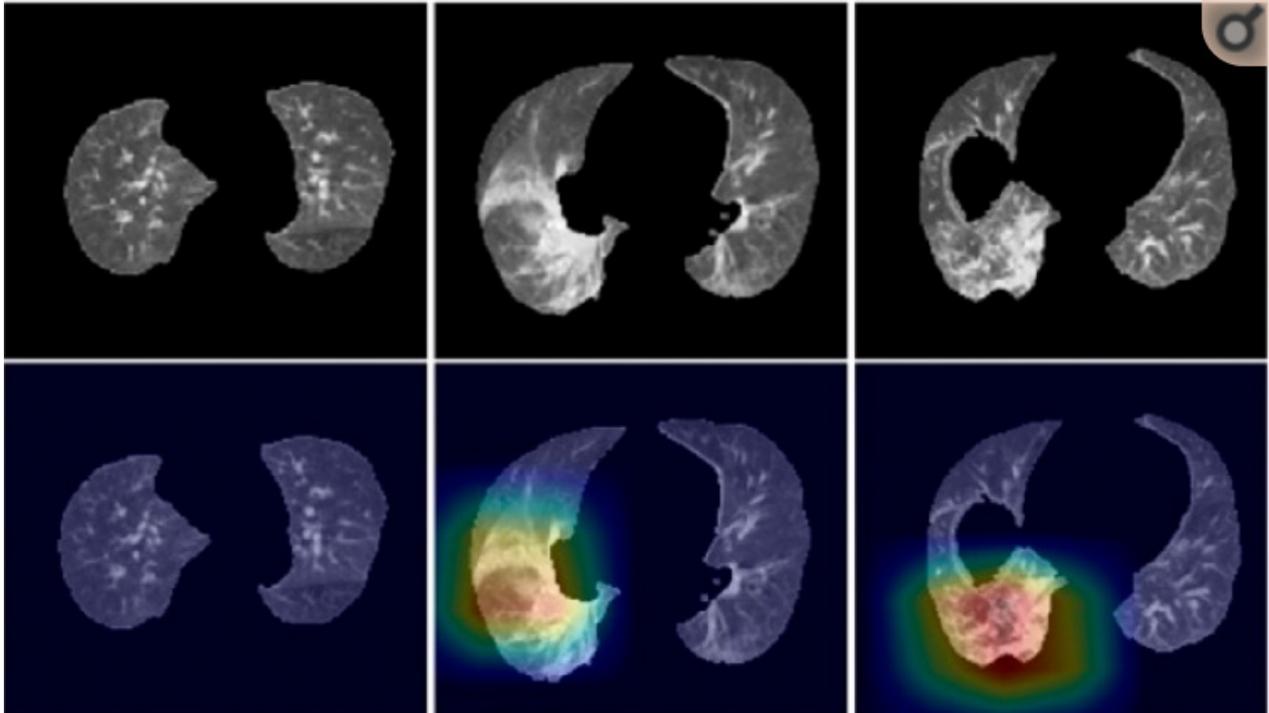
En esta búsqueda se identificaron 2 340 registros, de que 1 230 se examinaron luego de eliminar duplicados. Se excluyeron 1 065 artículos por no cumplir con los criterios de inclusión; seguido a esto, se evaluó la elegibilidad de 165 artículos de texto completo y se descartaron 157 al escanear dicho texto; de resultado se incluyeron 10 estudios en la revisión sistemática. La calidad de las investigaciones empleadas fue evaluada por los revisores y se ajustó a la versión revisada de

Figura 3. Ejemplo 1 representativos de los mapas de calor de atención generados con el método Grad-CAM. Este es un sistema de IA que usa una capa convolucional final para producir un mapa de localización aproximado que resalta lo importante en la imagen y predice el concepto, en este caso, la neumonía. El color rojo resalta la región de activación asociada con la clase predicha. La imagen superior muestra la neumonía por COVID-19 y la imagen inferior la adquirida en la comunidad.



Fuente: Tomada para fines académicos de: *artificial intelligence distinguishes COVID-19 from community acquired pneumonia on Chest CT* (38).

Figura 4. Ejemplo 2 representativo de los mapas de calor de atención generados con el método Grad-CAM. Este es un sistema de IA que usa una capa convolucional final para producir un mapa de localización aproximado que resalta lo importante en la imagen y predice el concepto, en este caso, la neumonía. El color rojo resalta la región de activación asociada con la clase predicha. La imagen superior muestra la neumonía por COVID-19 y la imagen inferior la adquirida en la comunidad.



Fuente: Tomada para fines académicos de: *artificial intelligence distinguishes COVID-19 from community acquired pneumonia on Chest CT* (38).

Quality Assessment of Diagnostic Accuracy Studies (QUADAS).

En los estudios se analizaron los indicadores de sensibilidad, especificidad, AUC y F1, y las curvas de características operativas del receptor (ROC) para evaluar la precisión del modelo de IA. En dicho proceso, las cifras ROC proporcionaron sensibilidad y especificidad promedio en los estudios incluidos, esto con un intervalo de confianza (IC) del 95 % del punto de operación resumido.

Las 10 investigaciones fueron estudios de cohorte y se resalta que no se encontraron ensayos clínicos aleatorizados que cumplieran con los criterios de inclusión, así, no se incluyeron otras revisiones sistemáticas. En estos se buscaron medidas de sensibilidad, especificidad y AUC según las curvas ROC y F1, por consiguiente, todos utilizaron datos retrospectivos y no fueron de acceso abierto. Además, se destaca que siete estudios excluyeron imágenes de baja calidad, y la comparación entre los modelos de IA y los profesionales sanitarios, mediante un igual conjunto de pruebas, solo proporcionó 6 estudios, por tanto, se ultimó que todos fueron ejecutados en países del sudeste asiático.

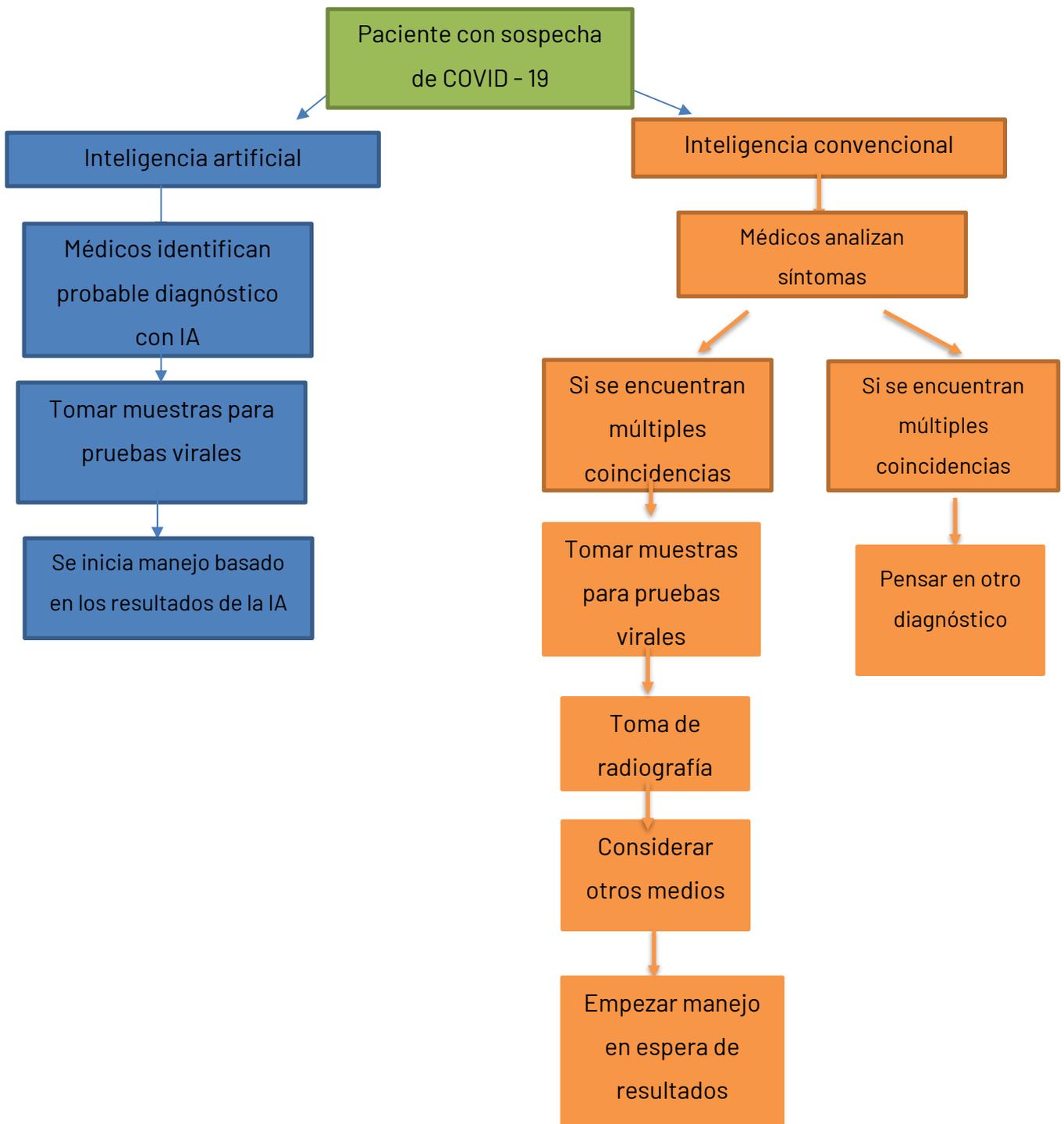
En los 10 estudios se impulsaron sistemas capaces de reconocer y extraer características representativas locales 2D y 3D globales. Dichos sistemas usados tomaron una serie de cortes de TC como entrada

y generaron particularidades para los cortes correspondientes; estas se combinaron por medio de una operación de agrupación máxima. De esta manera, el mapa de características final se alimentó de una capa completamente conectada y de una función de activación para producir una puntuación de probabilidad para cada tipo de neumonía utilizada. En este aspecto, los datos de programación fueron variados, algunos de los conjuntos incluyeron imágenes, pero no datos auxiliares como anotaciones, antecedentes demográficos y resultados de laboratorio. Algunas de las anotaciones fueron básicas, como la neumonía por COVID-19 o la no existencia de esta; otras fueron más completas, incluido el etiquetado anatómico o los resultados de las pruebas de diagnóstico de anticuerpos.

Entre dichos estudios se destacó el de (Bai et al., 2020), estos implementaron el sistema *EfficientNet B4* diagnóstico de neumonía de tomografías computarizadas (38). En dicho trabajo se utilizó el diagnóstico de los seis radiólogos sobre los pacientes correspondientes, con el objetivo de evaluar la eficiencia de los resultados obtenidos por un modelo de IA (38). El modelo de IA de estos expertos logró un 96 % de precisión, mientras que en la precisión promedio del diagnóstico de los radiólogos se obtuvo en un 85 %.

Por su parte, (Ying et al. 2020) diseñaron una red denominada *DRE-Net*, esta se basó en un sistema previamente entrenado y los resultados en su estudio

Figura 5. Procedimiento general de aplicaciones basadas en IA y no basadas en IA que ayudan a los médicos generales a identificar los síntomas del COVID-19(33).



Fuente: Tomado para fines académicos de Artificial Intelligence (AI) Applications for COVID-19 Pandemic.vaishya R,Javaid M,Khan IH, Haleem A.[internet].2020

mostraron que superó a otros modelos al lograr 0.92 y 0.95 de las puntuaciones AUC para los niveles de imagen y humanos (39).

Otro de los trabajos fue el de (Li et al.), en este se propuso una *COVNet* que diferenciaba la neumonía por COVID-19, la no neumonía y la neumonía adquirida en la comunidad (38). En su estudio se consideraron 4 352 tomografías computarizadas de tórax de 3 322 pacientes y se aplicó una operación de agrupación

máxima a las características obtenidas de *COVNet* (38). Cabe resaltar que este estudio no comparó el rendimiento de sus sistemas con radiólogos expertos (38). Con similares ideas, (Chen et al.) programaron un modelo que logró una precisión por paciente del 95.24 % y una precisión por imagen del 98.85 % en un conjunto de datos retrospectivos internos (40); en el conjunto de datos externos se alcanzó una precisión del 96 % (40). Con la ayuda del modelo, el tiempo de lectura de los radiólogos se redujo considerablemente en un 65 % (40)

y la sensibilidad y la especificidad fueron superior al 96 % (40).

Asimismo, en la investigación de (Jin C et al.), el sistema basado en una red neuronal convolucional profunda fue capaz de obtener un área bajo la curva característica de funcionamiento del receptor (AUC) del 97.81 %, esto para la clasificación multidireccional en una cohorte de prueba de 3 199 exploraciones, AUC de 92.99 % y 93.25 % en dos conjuntos de datos disponibles públicamente, CC-CCII y *MosMedData* respectivamente (41). En una indagación de lectores que involucró a cinco radiólogos, el sistema de IA superó a todos en tareas más desafiantes y a una velocidad de dos órdenes de magnitud por encima de ellos (41).

En igual línea, (Mei et al.) probaron un sistema de IA que alcanzó un área bajo la curva de 0.92 y tuvo similar sensibilidad en comparación con un radiólogo torácico experimentado (42). El sistema de IA también mejoró la detección de pacientes positivos para COVID-19 por intermedio de RT-PCR que presentó TC normales. En este orden de ideas, se identificó, correctamente, a 17 de 25 pacientes (68 %), y los radiólogos los clasificaron a todos como Covid-19 negativos (42). Con parecidos argumentos, en el estudio de (Ni et al.) el algoritmo mostró una sensibilidad superior de 1.00 (intervalo de confianza (IC) del 95 %: 0.95, 1.00) y una puntuación F1 de 0.97 en la detección de lesiones a partir de imágenes de TC de pacientes con neumonía COVID-19.

Mientras que, para la base del lóbulo por pulmón, el algoritmo logró una sensibilidad de 0.96 (IC del 95 %: 0.94; 0.98) y una puntuación F1 ligeramente inferior de 0.86. El algoritmo fue mucho más rápido que los residentes en la evaluación de las imágenes de TC (todas $p < 0,017$) (43).

Discusión

La pandemia debido al COVID-19 ha sido el espacio propicio para estudiar diversas tecnologías (9) y la oportunidad de verificar el desempeño de la IA, esta ha sido probada en diversos estudios en el contexto del COVID-19 y ha demostrado un rendimiento equivalente o incluso mejor para los profesionales de la salud, en términos de sensibilidad y especificidad (5). No obstante, se necesitan estándares de informes rigurosos con validación externa y comparación con los profesionales de la salud para la aplicación y estandarización de la IA en el campo médico.

De igual manera, en esta revisión sistemática se encontró que los algoritmos de IA pueden usarse para el diagnóstico de neumonía por COVID-19, esto a partir de material de imágenes de radiología médica con un rendimiento equivalente al de los profesionales de la salud. Dicho rendimiento de los sistemas de IA evaluados no solo se demuestra con los porcentajes de sensibilidad y especificidad, pues el área bajo la curva AUC de la curva ROC es un parámetro para valorar

la bondad de una prueba diagnóstica que produce resultados continuos y se interpreta en la probabilidad de que dos individuos, uno sano y otro enfermo, se clasifiquen correctamente. El AUC de los 5 estudios medidos fue superior a 0.9, esto indicó una sensibilidad y especificidad conjuntas óptimas.

La sensibilidad y especificidad de IA se ha validado con suficiencia en estudios previos y con otras condiciones como las metástasis tumorales (2,10). Por consiguiente, el excelente rendimiento de IA en la identificación de imágenes con alta velocidad, alta precisión y reducción significativa causa gran expectativa. En relación con esto, (Liu et al.) llevaron a cabo una revisión sistemática y un metanálisis, encontraron que el rendimiento diagnóstico de los modelos de aprendizaje profundo, a partir de imágenes médicas, era equivalente al de los profesionales de la salud en la clasificación de enfermedades, con una sensibilidad del 87.0 % y una especificidad del 92.5 %, esto funcionó de sustento para el uso clínico de los modelos de aprendizaje profundo (44).

Simultáneamente, un reciente metaanálisis de (Zheng et al.) demostró que, para condiciones como tumores cerebrales metastásicos, la IA mostraba una sensibilidad combinada del 82 % (IC 95 %, 79 % - 84 %), una especificidad del 84 % (82 % - 87%) y AUC de 0.90 (0.87 - 0.92) (2). Añadido a esto, el análisis de diferentes algoritmos de IA expuso una sensibilidad combinada del 87 % (83 % a 90 %) para el aprendizaje automático, un 86 % (82 % a 89 %) para el aprendizaje profundo, y una especificidad combinada del 89 % (82 % a 93 %) para el aprendizaje automático y 87 % (82 % - 91 %) para aprendizaje profundo (2).

En la presente investigación se observó una heterogeneidad, estadísticamente significativa, entre los estudios incluidos secundarios a varios defectos metodológicos comunes. Se evidenció que solo 6 de los 10 estudios encontrados utilizaron comparación con un radiólogo experto (1,35,41,43,45,46). Es por ello por lo que evaluar el rendimiento de la IA sin compararlo con su contraparte humana hace a los resultados no confiables cuando se aplican en el entorno clínico. Afortunadamente, los 6 estudios emplearon, en su comparación, un igual conjunto de datos de prueba.

Otra falencia identificada fue que los estudios reclutaron radiólogos con diferentes años de experiencia y distintos números, y algunos no los capacitaron previamente; todo lo anterior puede influir en el resultado. Además, solo uno de los estudios manejó otros materiales clínicos: la historia clínica electrónica y la información clínica de los pacientes (42). En efecto, unos emplearon un conjunto de datos de acceso abierto en línea, en lugar de ejecutarlo en el entorno clínico real. Incluso, se encontró otra debilidad en cuanto al estándar para diagnosticar neumonía, pues unos estudios solo recurrieron a la opinión de un único radiólogo como sustento, lo que pudo no ser

convinciente (17).

Desde diferente ángulo, otro de los hallazgos fue el rendimiento de las redes preformadas en situación de urgencia diagnóstica, esto motivado a que la mayoría de los sistemas desarrollados se modelaron con las mejoras en las redes preentrenadas. Esto permitió optimizar la precisión diagnóstica del COVID-19 en imágenes de TC; las redes preformadas se pueden utilizar con mucho éxito en todas las etapas de la clasificación de imágenes. En suma, los investigadores señalaron que muchos de los sistemas diseñados se modelaron con modificaciones de dichas redes para favorecer la precisión de clasificación del COVID-19 en imágenes de TC después de las etapas de preprocesamiento y segmentación.

Algunos investigadores clasificaron los casos de COVID-19 con técnicas de aprendizaje automático en lugar de aprendizaje profundo, esto alcanzó resultados de alto reconocimiento y permitió obtener un buen desempeño en cortos periodos de tiempo. La velocidad conferida por esta práctica es ideal en un contexto de urgencia como el que plantea esta enfermedad en una sala de urgencias o una unidad de cuidado intensivo.

Un aspecto negativo fue que, desafortunadamente, no manipularon indicadores de rendimiento diagnóstico uniformes, debido a que la mayoría usaron indicios como la sensibilidad, la especificidad y el AUC. Cabe mencionar que, cuando se informó el número de pacientes en los grupos intervención en el conjunto de datos de prueba, la sensibilidad y la especificidad se pudieron usar para la construcción de tablas de contingencia. De igual manera, fue poco alentado que indicadores como la precisión, la proporción de datos, la puntuación F1 y la memoria, comunes en el campo de la informática, no fueron medidos y ninguno de los estudios contó con validación externa. Para mejorar esto, se necesita con urgencia una estricta validación externa en el desarrollo del modelo de diagnóstico.

Conclusiones

A partir de la revisión sistemática de la literatura, y en respuesta a los objetivos planteados en este estudio, se llegó a las siguientes conclusiones.

El diagnóstico de neumonía por COVID-19, por intermedio de algoritmos de IA, tiene un gran potencial (47), en este sentido, con prudencia es posible indicar que estos pueden usarse para el diagnóstico de la enfermedad con base en de imágenes de radiología médica y con un rendimiento equivalente o mejor para los profesionales de la salud en términos de sensibilidad y especificidad, esto proporciona una base para su aplicación clínica (38).

Como un corolario más, la aplicación clínica de IA puede aliviar la escasez de recursos médicos, mejorar la tasa de detección, la precisión y el pronóstico de los pacientes (1). También se pudo concluir que los retos en personal y tiempo, impuestos por la pandemia, hacen

necesario el desarrollo de herramientas diagnósticas capaces de contar con la agilidad, la precisión y la exactitud suficientes para salvar vidas humanas y retrasar el avance del virus (48). Por último, se acota que se necesitan más investigaciones de alta calidad sobre la aplicación de la IA en el campo médico, no obstante, se notó, con mucha expectativa, que el potencial de la IA para la adaptación de la práctica clínica es sólida, confiable y prometedora (49 - 51).

Bibliografía

1. Xiong Z, Wang R, Bai H, Halsey K, Mei J, Li Y. *Artificial Intelligence augmentation of radiologist performance in distinguishing COVID-19 from pneumonia of other origin at Chest CT*. *Radiology*. 2020; 296(3):3.
2. Q Z, L Y, B Z, J L, K G, Y L. *Artificial intelligence performance in detecting tumor metastasis from medical radiology imaging: A systematic review and meta-analysis*. *EClinicalMedicine*. 2020; 31.
3. Mei X, Lee H, Diao K, Huang M, Lin B, Liu C. *Artificial intelligence-enabled rapid diagnosis of patients with COVID-19*. *Nat Med*. 2021; 26(8).
4. Adly A, Adly A, Adly M. *Approaches based on artificial intelligence and the internet of intelligent things to prevent the spread of COVID-19: scoping review*. *Journal of medical internet research*. 2020; 22.
5. Hu Z, Ge Q, Li S, Jin L, Xiong M. *Artificial intelligence forecasting of Covid-19 in China*. *arXiv*. 2020.
6. Shi F, Wang J, Shi J, Wu Z, Wang Q, Tang Z. *Review of artificial intelligence techniques in imaging data acquisition, segmentation, and diagnosis for Covid-19*. *IEEE Reviews in Biomedical Engineering*. 2020; 14.
7. Neri E, Miele V, Coppola F, Grassi R. *Use of CT and artificial intelligence in suspected or Covid-19 positive patients: statement of the Italian Society of Medical and Interventional Radiology*. *Radiol Medica*. 2020; 125(5).
8. Fan L, Li D, Xue H, Zhang L, Liu Z, Zhang B. *Progress and prospect on imaging diagnosis of Covid-19*. *Chinese J Acad Radiol*. 2020; 3(1).
9. D D, Z T, S W, H H, L G, Y L. *The role of imaging in the detection and management of Covid-19: a review*. *IEEE Rev Biomed Eng*. 2021; 14.
10. Oliver C, Altemus M, Westerhof T, Cheriyan H, Cheng X, Dziubinski M. *A platform for artificial intelligence based identification of the extravasation potential of cancer cells into the brain metastatic niche*. *Lab Chip*. 2021; 19(7).
11. Hosny A, Parmar CQJSL, Aerts H. *Artificial intelligence in radiology*. *Nature Publishing Group*. 2018; 18.
12. Miotto R, Wang F, Wang S, Jiang X, Dudley J. *Deep learning for healthcare: review, opportunities and challenges*. *Brief Bioinform*. 2017; 19(6).
13. Tsehay Y, Lay N, Roth H, Wang X, Kwak J, Turkbey B. *Convolutional neural network based deep-learning architecture for prostate cancer detection on multiparametric magnetic resonance images: medical imaging*; 2017.
14. Syed A, Zoga A. *Artificial intelligence in radiology: current technology and future directions*. *Thieme medical publishers, Inc*. 2018; 22.
15. Thrall J, Li X, Li Q, Cruz C, Do S, Dreyer K. *Artificial intelligence and machine learning in radiology: opportunities, challenges, pitfalls, and criteria for success*. *J Am Coll Radiol*. 2018; 15(3).
16. Thompson R, Valdes G, Fuller C, Carpenter C, Morin O, Aneja S. *Artificial intelligence in radiation oncology: a*

- specialty-wide disruptive transformation. *Radiotherapy and Oncology*. 2018; 129.
17. Kobayashi Y, Ishibashi M, Kobayashi H. How will "democratization of artificial intelligence" change the future of radiologists. *Japanese Journal of Radiology*. 2019; 37.
 18. Dumić-Čule I, Orešković T, Brkljačić B, Tiljak M, Orešković S. The importance of introducing artificial intelligence to the medical curriculum – assessing practitioners' perspectives. *Croatian Medical Journal*. 2020; 61.
 19. Mascarenhas V, Caetano A, Dantas P, Rego P. Advances in FAI Imaging: a focused review. *Curr Rev Musculoskelet Med*. 2020; 13(5).
 20. Rubin D. Artificial Intelligence in Imaging: the radiologist's role. *J am coll radiol*. 2019; 16(9).
 21. Gampala S, Vankeshwaram V, Gadula S ¿Is artificial intelligence the new friend for radiologists? A review article. *Cureus*. 2020; 12(10).
 22. E.H. A-A. Genetic variations among SARS-CoV-2 strains isolated in China. *Gene Reports*. 2020; 21.
 23. Bai YX, Xu YH, Wang X, Sun C, Guo Y, Qiu S. Advances in SARS-CoV-2: a systematic review. *Rev Med Pharmacol Sci*. 2020; 14(17).
 24. Chan J, Kwok K, Ma J, Wong Y. *Radiology and Covid-19*. Hong Kong Medical Journal. 2020; 26.
 25. Jacobi A, Chung M, Bernheim A, Eber C. Portable chest X-ray in coronavirus disease-19 (Covid-19): A pictorial review. *Clinical Imaging*. 2020; 64.
 26. Oterino et al. Pediatric chest x-ray in Covid-19 infection. *Eur J Radiol*. 2020.
 27. Salameh J, Leeflang M, Hooft L, Islam N, McGrath T. Thoracic imaging tests for the diagnosis of Covid-19. *Cochrane Database Syst Rev*. 2020; 9.
 28. Lim M, Pranata R. Coronavirus disease 2019 (Covid-19) markedly increased mortality in patients with hip fracture – A systematic review and meta-analysis. *J Clin Orthop trauma*. 2020.
 29. Xu X, Yu C, Qu J, Zhang L, Jiang S, Huang D. Imaging and clinical features of patients with 2019 novel Coronavirus SARS-CoV-2. *Eur J Nucl Med Mol Imaging*. 2020; 47(5).
 30. Patel N. Pediatric Covid-19: systematic review of the literature. *Am J Otolaryngol*. 2020; 41(5).
 31. Güneyli S, Atçeken Z, Doğan H, Altınmakas E, Atasoy K. Radiological approach to Covid-19 pneumonia with an emphasis on chest CT. *Diagnostic and Interventional Radiology*. 2020; 26.
 32. Ai T, Yang Z, Hou , Zhan C, Chen C, Lv W. Correlation of chest CT and RT-PCR testing for Coronavirus disease 2019 (Covid-19) in China: a report of 1014 Cases. *Radiology*. 2020; 296(2).
 33. Pan F, Ye T, Sun P, Gui S, Liang B, Li L. Time course of lung changes at chest CT during recovery from Coronavirus disease 2019 (Covid-19). *Radiology*. 2020; 295(3).
 34. Vaishya R, Javaid M, Khan I, Haleem A. Artificial intelligence (AI) applications for Covid-19 pandemic. *Diabetes metab syndr clin Res Rev*. 2020; 14(4).
 35. Li Z, Zhong Z, Li Y, Zhang T, Gao L, Jin D. From community-acquired pneumonia to Covid-19: a deep learning-based method for quantitative analysis of Covid-19 on thick-section CT scans. *Eur Radiol*. 2020; 30(12).
 36. Ito R, Iwano S, Naganawa S. A review on the use of artificial intelligence for medical imaging of the lungs of patients with coronavirus disease 2019. *Diagnostic and Interventional Radiology*. 2020; 26.
 37. Nagpal P, Narayanasamy S, Vidholia A, Guo J, Shin K, Lee C. Imaging of Covid-19 pneumonia: patterns, pathogenesis, and advances. *The british journal of radiology*. 2020; 93.
 38. Li L, Qin L, Xu Z, Yin Y, Wang X, Kong B. Using artificial intelligence to detect Covid-19 and community-acquired pneumonia based on pulmonary ct: evaluation of the diagnostic accuracy. *Radiology*. 2020; 196(2).
 39. Ying S, Zheng S, Li L, Zhang X, Zhang X, Huang Z. Deep learning enables accurate diagnosis of novel coronavirus (Covid-19) with CT images. *medRxiv*. 2020.
 40. Chen J, Wu L, Zhang J, Zhang L, Gong D, Zhao Y. Deep learning-based model for detecting 2019 novel Coronavirus pneumonia on high-resolution computed tomography. *Sci Rep*. 2020; 10(1).
 41. Jin C, Chen W, Cao Y, Xu Z, Tan Z, Zhang X. Development and evaluation of an artificial intelligence system for Covid-19 diagnosis. *Nat Commun*. 2020; 11(1).
 42. Mei X, Lee H, Diao K, Huang M, Lin B, Liu C. artificial intelligence-enabled rapid diagnosis of patients with Covid-19. *Nat Med*. 2020 26; 8.
 43. Ni Q, Sun Z, Qi L, Chen W, Yang Y, Wang L. A deep learning approach to characterize 2019 coronavirus disease (Covid-19) pneumonia in chest CT images. *Eur Radiol*. 2020; 30(12).
 44. P J, X J, W K, Y L, H L, F M. Artificial intelligence in gastric cancer: a systematic review. *Journal of Cancer Research and Clinical Oncology*. 2020; 146.
 45. Jin C, Chen W, Cao Y, Xu Z, Tan Z, Zhang X. Development and evaluation of an artificial intelligence system for Covid-19 diagnosis. *Nat Commun*. 2020; 11(1).
 46. Suri J, Agarwal S, Gupta S, Puvvula A, Biswas M, Saba L. A narrative review on characterization of acute respiratory distress syndrome in Covid-19-infected lungs using artificial intelligence. *Comput Biol Med*. 2020; 130.
 47. Alsharif W, Qurashi A. Effectiveness of Covid-19 diagnosis and management tools: a review. *Radiography*. 2020; 21.
 48. Brown N. A novel peer-to-peer contact tracking application for Covid-19 and future pandemics. *Diabetes metab syndr Clin Res Rev*. 2020; 14(5).
 49. R V, M J, IH K, A H. Artificial intelligence (AI) applications for Covid-19 pandemic. *diabetes metab syndr clin res rev*. 2020; 14(4).
 50. Suri J, Puvvula A, Biswas M, Majhail M, Saba L, Faa G. Covid-19 pathways for brain and heart injury in comorbidity patients: a role of medical imaging and artificial intelligence-based Covid- 19 severity classification: a review. *Computers in Biology and Medicine*. 2020; 124.
 51. Jalaber C, Lapotre T, Morcet-Delattre T, Ribet F, Jouneau S, Lederlin M. Chest CT in Covid-19 pneumonia: a review of current knowledge. *Diagnostic and Interventional Imaging*. 2020; 101. x